

Generalized Hill-Stability Criteria for Hierarchical Three-Body Systems at Arbitrary Inclinations

Evgeni Grishin, Hagai B. Perets, Yossef Zenati and Erez Michaely
Physics Department, Technion - Israel Institute of Technology, Haifa, Israel 3200003

A fundamental aspect of the three-body problem is the stability of triple systems. Most stability studies have focused on the co-planar three-body problem, deriving analytic criteria for the dynamical stability of such pro/retrograde systems. Numerical studies of inclined systems phenomenologically mapped their stability regions, but did not explain their physical origin, nor provided satisfactory fit for the dependence of stability on the mutual inclinations. Here we present a novel approach to study the stability of hierarchical three-body systems at arbitrary inclinations. This approach accounts not only for the instantaneous stability of such systems, but also for the secular stability and evolution through Lidov-Kozai cycles and evection. Thereby we are able to generalize the Hill-stability criteria to arbitrarily inclined triple systems, and explain the existence of quasi-stable regimes and characterize the inclination dependence of their stability. We complement the analytic treatment with an extensive numerical study, to test our analytic results. We find excellent correspondence up to high inclinations ($\sim 120^\circ$), beyond which the agreement of our analytically derived stability criteria becomes marginal. At such high inclinations the stability radius is larger, the ratio between the outer and inner period of the triple systems becomes comparable, and our secular averaging approach is no-longer strictly valid. We therefore combine our analytic results with fits to the numerical results at high inclinations to obtain a generalized stability formula for triple systems at arbitrary inclinations. Besides providing a generalized secular-based physical explanation for the stability at non co-planar systems, our results have direct implications for any triple systems, and in particular binary planets and moon/satellite systems; we briefly discuss the latter as a test case for our models.

I. INTRODUCTION

The three-body problem is an old topic in celestial mechanics, with wide astrophysical applications in the Solar system and beyond [1–3]. Hierarchical triple systems are systems in which an inner binary orbits a more distant object on an outer orbit, with some mutual given inclination between the inner and outer orbits. A fundamental aspect of a hierarchical triple system is its stability. A system is considered stable if no collision or escape of one of the bodies occurs after a large number of orbital periods. The natural length scale of stability is the Hill radius $r_H = a_{out}(\mu/3)^{1/3}$, where a_{out} is the distance from the distant outer body with mass m_{out} to the center of mass of the binary and $\mu \equiv (m_1 + m_2)/m_{out}$ is the binary to perturber mass ratio. If the inner binary semi-major axis (SMA) is larger than the Hill radius $a_{in} > r_H$, the binary is unstable, since tidal forces from the perturber shear apart the binary. Conversely, if $a_{in} \ll r_H$, then the binary is stable and the perturbations from m_{out} are small. Early studies of *co-planar* triple systems [4–6] have shown that prograde orbits are stable for $a_{in} \approx 0.5r_H$ while retrograde orbits are stable for twice the distance.

The first analytic study to explore the critical stability radius at *arbitrary* inclination was done by Innanen [7, 8]. He finds that the critical stability radius in an *increasing* function of the inclination, thus the most stable orbits are retrograde, consistent previous results of co-planar orbits. However, numerical simulations show that the critical radius is compatible with the analytical expectation only for moderately inclined orbits. For highly

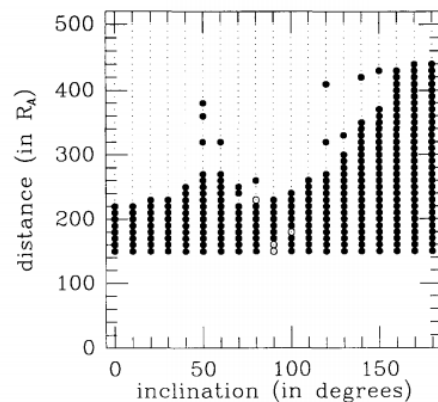


FIG. 1. The critical distance, which divides stable from unstable orbits, as a function of initial inclination. Large solid dots signifies stable orbit for at least 5 asteroid years. Small dots are unstable, escaping orbits. Open circles are collisional orbits. Originally shown by [6] (their Fig. 15). Their units are $r_H \approx 450R_A$.

inclined orbits, the critical radius starts to decrease at $\sim 60^\circ$, and increase again only at higher inclinations, forming form a bowl-like shape (see Fig. 1; originally shown by [6], their units are $r_H \approx 450R_A$).

The general equations of motion of the three-body problem cannot be solved analytically [9], however secular averaging analysis can be used to describe a wide range of cases. For hierarchical systems, pioneering works of Lidov [10] and Kozai [11] have shown that the torque of the outer binary can induce significant quasi-periodic oscillations in the inclination and eccentricity of the in-

ner binary over long, secular timescales, under certain configurations.

The Lidov-Kozai (LK) mechanism is obtained by double averaging (DA) over the orbits. The averaging process does not take into account short period terms that affect the secular evolution. Ćuk and Burns [12] considered corrections due to evection and Lunar theory for circular orbits in the context of irregular moons around giant planets. Recently, a critical analysis of the DA technique and its correction, as well as generalization of [12] for arbitrary inclination and eccentricity was studied by [13].

In previous analytic studies of Hill stability, the impact of secular evolution, and in particular Lidov-Kozai cycles were not fully addressed. Numerical studies provided insights and phenomenological mapping for the stability criteria of such systems [6, 14, 15], but did not explain their physical origin, nor provided analytic derivations and satisfactory fits for the dependence of stability on the mutual inclinations. In this paper we account for secular evolution in exploring the stability of triple systems (namely the LK mechanism and evection), and provide a physical understanding of its behavior. In addition, we perform an extensive numerical three-body study of the stability criteria, confirming our analytic solutions (finding excellent correspondence up to high inclinations) and complement them in regimes where the secular averaging approach we use is no-longer valid.

The paper is organized as follows. In Sec. II we briefly provide the background for previous analytic study of the instantaneous Hill stability at arbitrary inclinations based on [8]. In Sec. III We first describe the LK mechanism, and then couple it to stability analysis. We then show how secular evolution affects the results, explaining the discrepancy between [8] and [6]. In Sec. IV we explore how evection further improves the stability analysis. In Sec. V, we numerically integrate the three-body problem to confirm and complement our analytic approach, providing a generally useful fitting formula for triple stability at arbitrary inclinations. Finally we discuss our results and their implications and summarize in Sec. VI.

II. INSTANTANEOUS HILL STABILITY CRITERIA AT ARBITRARY INCLINATIONS: CORIOLIS ASYMMETRY

Before deriving a generalized stability criteria which accounts for *secular* processes, let us first review the main results of Innanen [8]'s work on the Coriolis asymmetry, and derive an expression for the *instantaneous* stability criteria in a modern fashion. This expression, as remarked above, fail to correctly capture the observed stability behavior, however it is not incorrect by itself, but rather has to be coupled to secular effects, as we show in the next sections.

Consider the restricted, mass hierarchical three-body system. We will use "star" of mass m_{out} , "planet" of

mass m_1 and "satellite" of mass m_2 for the hierarchical masses $m_{out} \gg m_1 \gg m_2$, but alternatively, it could apply to other astrophysical systems. The planet revolves around the star at SMA a_{out} , and the satellite revolves around the planet at SMA $a_{in} \ll a_{out}$. For convenience, we omit the subscript "in" for the orbital elements of the inner binary (e.g. the assumption on hierarchical system is $a \ll a_{out}$).

Consider a binary with arbitrary inclination. The instantaneous distances of the satellite from the star and the planet are \mathbf{R} and \mathbf{r} respectively. The acceleration of the satellite is given by [8]

$$\mathbf{a} = \ddot{\mathbf{R}} + \dot{\boldsymbol{\Omega}}_{out} \times \mathbf{r} + \boldsymbol{\Omega}_{out} \times (\boldsymbol{\Omega}_{out} \times \mathbf{r}) + 2\boldsymbol{\Omega}_{out} \times \mathbf{v}_r, \quad (1)$$

where $\boldsymbol{\Omega}_{out}$ is the angular frequency vector around the star, $\mathbf{v}_r = \omega_1 \mathbf{r}$ is the velocity of the satellite and ω_1 is the angular frequency around the planet. The first term is $\ddot{\mathbf{R}} = \boldsymbol{\Omega}_{out} \times (\boldsymbol{\Omega}_{out} \times \mathbf{R})$, the gravitational acceleration from the star, the second term is the change of angular frequency, the third term is the gravitational acceleration from the planet, and the last term is $2\boldsymbol{\Omega}_{out} \omega_1 \mathbf{r} \cos i$, the Coriolis term, where i is the inclination angle of the binary. For constant $\boldsymbol{\Omega}_{out}$, the second term vanishes. At the critical radius, tidal forces from the Sun and other forces equal the gravitational pull of the planet. Thus, at the critical radius, the acceleration is zero, and Eq. (1) is an algebraic equation for the minimal angular frequency (or maximal radius)

$$\omega_1^2 - 2\omega_1 \Omega_{out} \cos i - 3\Omega_{out}^2 = 0 \quad (2)$$

For prograde orbit ($i = 0$), $\omega_1 = 3\Omega_{out}$ and the critical radius is $r_c = r_H / 3^{1/3} \approx 0.7r_H$. For retrograde orbit ($i = 180^\circ$), $\omega_1 = \Omega_{out}$ and the critical radius is $r_c = 3^{1/3}r_H = 1.44r_H$. Thus, the Coriolis force is a stabilizing force for retrograde orbits and destabilizing force for prograde orbits.

It is possible to find the limiting radius for any inclination. [8] finds the ratio of the prograde to retrograde orbit for each inclination. Alternatively, we solve the quadratic equation (2) and find the limiting radii explicitly. The only positive solution to Eq. (2) is $\omega_1(i)/\Omega_{out} = g(i)$, where $g(i) \equiv \cos i + (3 + \cos^2 i)^{1/2}$. Note that the Coriolis asymmetry is evident in the first term, and that the solution reduces to $\omega_1 = 3\Omega_{out}$ for $i = 0$ and $\omega_1 = \Omega_{out}$ for $i = 180^\circ$. The limiting radius is then

$$r_c(i) = \left(\frac{Gm_1}{\omega_1^2(i)} \right)^{1/3} = 3^{1/3} r_H g(i)^{-2/3} \quad (3)$$

Note that $g(i)$ is a monotonic decreasing function for $i \in [0, \pi]$, thus $r_c(i)$ is an increasing function.

The limiting radii derived in (3) is obtained by taking the acceleration of the satellite to be zero. In reality, the binary could break apart before it reaches the limiting radius r_c . Indeed, numerical integrations (e.g. [6, 14, 16]) show that the transition from stable to unstable binary

is $\approx 0.5r_H$ for prograde orbits, and $\approx 1r_H$ for retrograde orbits. Thus, a fudge factor $q \leq 1$ can be introduced to effectively rescale the stability limit of Eq. (3). Note that for prograde orbit the stability limit is consistent with the critical Jacobi integral that allows escape [16].

In the latter discussion, one has to remember that the three body problem is chaotic, and the transition boundary from stable to unstable orbit is not strictly a line, but rather has a complex, fractal, structure. The final fate of the orbits near the stability limit sensitively depends on the initial conditions, and large number of orbits have to be integrated in order to get a statistical understanding of the stability limit. We explore the parameter space numerically in Sec. V.

The numerical integrations of [6] fit the expectations for almost co-planar inclinations (i.e. such that $\sin i$ is small). However, for high inclinations (i.e. such that $\sin i$ is large), the critical radius significantly deviates from the analytical expectations. In particular, the critical radius is *decreasing* for inclinations between $\sim 60^\circ$ and $\sim 90^\circ$, and is substantially lower than expected for inclinations up to $\sim 150^\circ$ (see Fig. 1).

III. NOVEL LIDOV-KOZAI-HILL SECULAR INSTABILITY CRITERIA AT ARBITRARY INCLINATIONS

In order to derive a generalized stability criteria which accounts for *secular* processes, we first review the secular LK mechanism. We then couple the stability criteria (3) to the secular LK mechanism to modify and derive a *generalized secular* stability criteria and compare it with a numerically integrated stability map. As we show, the novel Lidov-Kozai-Hill stability criteria qualitatively captures the instability properties at large inclinations. We further correct this criteria by accounting for evection (Lunar theory) processes in the next section.

A. Lidov-Kozai mechanism

The presence of a third, distant body in hierarchical system causes the orbital elements of the binary to change with time. The angles of ascending node Ω and argument of periape ω precess, while other orbital elements oscillate quasi-periodically with small magnitude, that depends of the ratio a/a_{out} . On average, the energy and angular momentum of the binary are conserved over orbital period of the distant body.

In the Lidov-Kozai (LK) resonance, the argument of periastron ω librates, rather than oscillates. Even for orbits not it resonance, the LK mechanism exchanges the binary eccentricity and inclination, obtained for large enough initial inclination i . For hierarchical triple systems the typical secular LK timescale is

$$T_{LK} \approx \frac{m_{out} + m_1 + m_2}{m_{out}} (1 - e_{out}^2)^{3/2} \frac{P_{out}^2}{P} \quad (4)$$

For the restricted three body problem and $e_{out} = 0$, it reduces to

$$T_{LK} \approx \sqrt{3} P_{out} \left(\frac{a}{r_H} \right)^{-3/2} = 3P \left(\frac{a}{r_H} \right)^{-3} \quad (5)$$

For most cases, the LK timescale is a *secular* timescale, much larger than the orbital period. Thus, the DA method is satisfactory. However, in our case, a is relatively comparable to r_H , thus, the LK timescale is not much larger than the binary period P , and consequently the system is only marginally (or quasi) secular. We discuss the implications and caveats of the quasi secular regime in Sec. VI

In the quadrupole DA approximation, the minimal inclination required for the LK mechanism is $i_c = \arcsin \sqrt{2/5}$, or $\approx 39.2^\circ$ for prograde orbits and 140.8° for retrograde orbits. In addition, the z component angular momentum of the binary, $L_z = \sqrt{1 - e^2} \cos i$ is constant. Thus, the maximal eccentricity attained during a KL cycle is

$$e_{max}(i, i_c) = \sqrt{1 - \frac{\cos^2 i}{\cos^2 i_c}} = \sqrt{1 - \frac{5}{3} \cos^2 i} \quad (6)$$

The larger the initial inclination i is, the larger is e_{max} . In addition, $e_{max} \rightarrow 1$ as the inclination approaches to 90° .

Using further octupole expansion beyond the quadrupole approximation, it was shown that under specific conditions even higher inner eccentricities can be obtained [17], and the inner orbit inclination can flip its orientation from prograde, to retrograde [18] (see [19] for a recent review and references therein of this eccentric LK (EKL) regime). Moreover, the secular averaging method itself fails at some point, and the systems can behave more erratically, leading to large eccentricity changes on dynamical timescales, and even more extreme eccentricities can be obtained [20]. Here we only take into account only the quadrupole term, other regimes are beyond the scope of this manuscript and will be discussed elsewhere.

B. Lidov-Kozai-Hill stability

Previous numerical simulations indicate that while the Coriolis asymmetry describes well co planar orbits, other effects besides the Coriolis asymmetry must be taken in to account for orbits with high inclination.

The original derivation of [8] is for circular orbits. In reality, secular perturbations due to the LK mechanism drives the binary to high eccentricities, that affect the limiting radii of stability. Given a maximal eccentricity e_{max} , the limiting radii is modified by the LK mechanism is

$$r_c^{(LK)}(i, e_{max}) \equiv r_c(i) \cdot f(e_{max}) \quad (7)$$

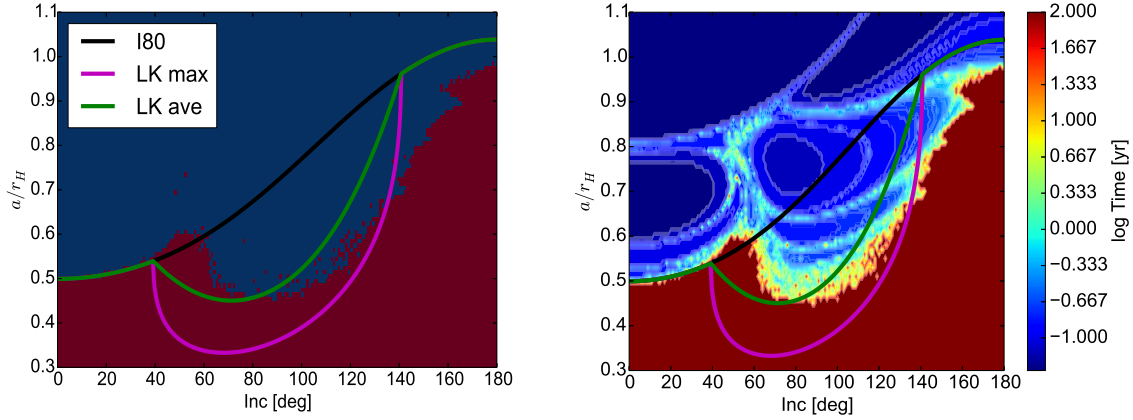


FIG. 2. Limiting radii versus numerical simulations of the stability map. Left: Blue pixel indicates stable orbit, red pixel is unstable orbit. The black line is the original curve given in Eq. 3 (normalized to $0.5r_H$ at zero inclination). The magenta line is the modified curve due to LK mechanism with f_{max} . The green line is the modified curve due to LK mechanism with $f_{ave}(e) = 1/(1 + 0.5e^2)$. Right: Break up times for the stability map. The stable region has time = 100yr.

Where $r_c(i)$ is given by Eqn. (3), $f(e_{max})$ is some function of e_{max} , such that $f(0) = 1$. For secular timescales, the critical radius can be satisfied at the apoapsis, or the maximal distance of the orbit, i.e. if $r_c^{(LK)}(i)(1 + e_{max}) = r_c(i)$, thus $r_c^{(LK)} = r_c(i)/(1 + e_{max})$, or $f = f_{max}(e_{max}) = 1/(1 + e_{max})$.

Perets and Naoz [21] first pointed out the importance of the LK secular evolution for the stability. Here we follow-up on this issue and consider the effects of secular evolution, couple them to the [8] approach and derive a generalized criteria which well captures the results from three-body simulations. In order to explore this analytically we first repeat the numerical experiment of [6] with different integrator and more refined resolution (as described in sec. V A). We obtain a highly resolved stability map, to be compared with the analytic results we derive.

The left panel of Fig. 2 shows this stability map. Red pixel indicates stable orbit for 100 yr, while blue pixel is unstable orbit. Generally, the boundary of stability is chaotic, but the chaotic regions are narrow and still a boundary line could be traced. The boundary increases with inclination, up to $\sim 60^\circ$, then it abruptly decreases, and start to slowly increase again for retrograde orbits, up to $\sim 1r_H$ for inclinations near 180° . The black curve is given in Eq. 3; it is increasing with increasing inclination, contrary to the boundary of the stability. The magenta curve is the correction due to LK mechanism, $r_c^{(LK)}$ with $f_{max} = (1 + e_{max})^{-1}$ in Eqn. 7 (i.e. “LK max”). The curve is bowl-shaped in the LK dominated region of high inclination, consistent with decrease, and subsequent increase of the stability boundary. We note that even though we have used a different integrating scheme and different masses and integration times, the results of the stability map reproduce the results of HB91.

The right panel of Fig. 2 shows the break up time of

the same simulated map. Dark red times are stable orbits with $t = 100$ yr. The shortest break up times are blue (comparable to dynamical times), while longer break up times are more red. We see that generally in the region where $i \geq 60^\circ$ the break up times are longer, indicating that the stability limit in this region is dominated by (quasi) secular effects. The additional structures of the break up timescale in the unstable region are probably due to additional resonances of the system, are out of the scope of this paper and will be studied elsewhere.

Though morphologically similar to the simulated stability boundary, the LK max curve has discrepancies both in the “turnoff point” where the LK mechanism is effective, and in the depth of the bowl.

A possible explanation to the discrepancy of the depth of the bowl is that near the stability limit, the timescales involved are not secular, but rather comparable to the dynamical timescale, i.e. the binary period P . In this case, the effective radius is not the apoapsis, but rather the average separation $\langle a \rangle_P = a(1 + 0.5e^2)$ [15]. The maximal average distance is found by setting $e = e_{max}$, thus the correction factor for taking the average distance is $f_{ave}(e_{max}) = 1/(1 + 0.5e_{max}^2)$. The green curve in Fig. 2 shows the correction due to LK mechanism, $r_c^{(LK)}$ with f_{ave} in Eqn. (7) (i.e. “LK ave”). It maintains the morphology of the shape of the stability boundary, though still is an effective shift in the critical inclinations of the LK mechanism.

We ran the same stability map for longer times of 10^4 yr, 100 times more that the previous integration time, and $\sim 10^3$ times more than the typical LK time-scale. Generally, the morphology of the stability map remains the same. The most notable difference is that some of the orbits in the chaotic region that were stable after 100 years are unstable after 10^4 years. Besides this (expected) difference, the picture remains the same. We

conclude that the stability grid is (almost everywhere) long lived, and assume hereafter that the same picture remains unaltered even for longer integration times. We conclude that the end time of 100 yrs fully captures the picture of the stability map, and almost no information is lost with the integration time of 100 years (e.g., the "bump" near 60° is real and not a numerical artifact of short integration times).

IV. EVECTION AND LUNAR THEORY CORRECTIONS

Taking into account the secular LK mechanism successfully improved the stability limit of inclined orbits with respect to numerical integrations, both in scale and shape. However, the discrepancy of the effective "turn off" point where the LK mechanism is effective remains. The critical inclination where the slope of the stability limit changes is near $\sim 60^\circ$ in the numerical stability map, while in the LK mechanism, the critical inclination angle is much lower ($i_c \approx 39.2^\circ$).

Interestingly, the inclination distribution of prograde irregular satellites in the Solar System has a cut-off near 60° . In addition, the entire range of inclinations of $60^\circ < i < 140^\circ$ is devoid of satellites [22, 23]. In order to tackle the problem, we consider additional perturbations to the system, namely the evection resonance (ER) and Lunar theory (LT).

ER is a periodic, rather than secular, perturbation to a satellite's orbit with a resonant angle $\lambda - \omega$ where λ is the mean longitude and ω is the argument of periapsis. The mean longitude is a fast angle that varies with the orbital time-scale. The averaging over the orbital period do not take into account the ER [12].

One of the consequences of ER is precession of the longitude of pericenter $\varpi = \omega + \Omega$ [31]. From secular theory, the precession rate $\dot{\varpi} = (3/4)\Omega_{out}^2/\omega_1$ has an error of $\sim 50\%$ from its observed value for the Earth-Moon system. Newton considered this discrepancy as a major flaw in his theory of gravitation. Later it has been realized that the large errors arise from evection terms that remain even after the averaging of the orbital motion [12, 13]. The strength of the evection is determined by a parameter

$$m = \Omega_{out}/\omega_1 \quad (8)$$

For the moon, it is about $\sim 1/13$.

Various works on LT [24–26] have derived the precession of the pericenter in a form of power series in m

$$\dot{\varpi} = \Omega_p \left(\frac{3}{4}m + \frac{225}{32}m^2 + \frac{4071}{128}m^3 + \dots \right) \quad (9)$$

The second term arises from ER, while higher order terms are calculated from LT [27]. For negative m , the orbit is retrograde and ϖ slowly precesses or even librates.

Using a development of the disturbing function to quadrupole order, [12] have shown that the additional

terms cause additional precession of the orbital elements. One of the important consequence of the ER and LT terms is the effective shift of the critical inclination for LK resonance. The equation for the precession of apsides is (Eq. (38) is [12])

$$\frac{d\omega}{d\tau} = h_0(i) + mh_1(i)\cos i + \sum_{k=2}^9 C_{k+1}m^k \cos^k i \quad (10)$$

where $h_0(i) = 2 - 5\sin^2 i$, $h_1(i) = (225/64)(2 - \sin^2 i) + (9/32)[3\sin^2 i - \cos(2i)]$, $\tau \equiv (1 - e_{out}^2)^{-3/2}m\Omega_{out}t$ is the dimensionless time, and the Coefficients C_{k+1} are given in the literature [32].

LK resonance occurs when $d\omega/d\tau = 0$, which determines the critical inclination. The $h_0(i)$ term is responsible for the classical LK resonance, occurs at $i_c = \arcsin \sqrt{2/5} \approx 39.2^\circ$. The $h_1(i)$ terms comes from the ER terms in the disturbing function. Higher order terms C_k come from LT.

Eq. (10) has been solved iteratively in [12]. However, for m larger than ~ 0.25 the iterative method diverges. Thus, we solve eqn. (10) numerically by substituting each pair (i_j, m_k) in $A_{jk} \equiv |d\omega/d\tau(i_j, m_k)|$, given in Eq. (10) and seeking the values of (i_j, m_k) for which $A_{jk} < \epsilon$ where ϵ is some small number. We took $\epsilon = 0.005$ such that each m has a solution $m(i)$.

Note that each m is coupled to $\cos i$, thus the retrograde orbit are accounted for with the convention that for retrograde orbit $m < 0$ and $i \rightarrow \pi - i$. In this convention, i is always smaller than 90° (see [12] for details and references therein).

Fig. 3 shows the numerical solution $m(i)$ such that $d\omega/d\tau(i, m(i)) = 0$. We see that for $m \neq 0$, the symmetry between the prograde and retrograde orbits breaks because of the $\cos i$ terms. We note that the slope and amplitude differ from Fig. (4) of [12], that solves the same equation. We have also checked the iterative method for low values of m and checked the fit against numerical integrations. Both methods are in good agreement with the numerical solution. The red dashed line is the linear fit

$$i(m) = i_{c,0} + i_{c,1}m, \quad (11)$$

where $i_{c,0} = 39.2^\circ$ (red point) and $i_{c,1} = 53.5^\circ$ is calculated analytically in appendix A.

In order to use an explicit formula for $i(m)$, we have used a least squares polynomial fit for the data of degree 40. The first few coefficients are

$$i(m) = 39.28 + 53.2m - 6.812m^2 + \mathcal{O}(m^3). \quad (12)$$

The full list of the coefficients is given in table II in Appendix B. Note that the first two coefficients in Eq. (12) are consistent with the analytical solution in Eq. (11). The bottom panel of Fig. 3 shows the deviation of the polynomial fit from the numerical solution, with RMS value of $\Delta i_{RMS} = 0.01^\circ$.

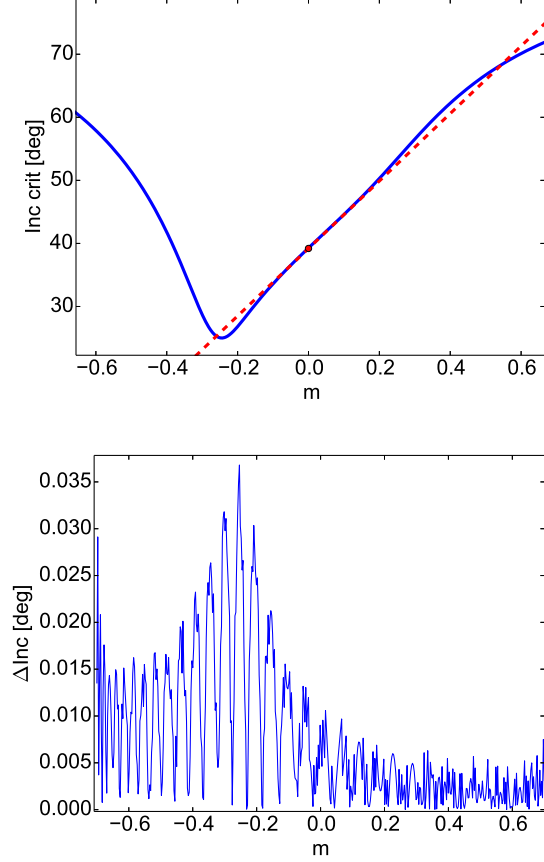


FIG. 3. Top: Numerical solution of Eq. (10) (blue). First order analytical solution of Eq. (10) (red, dashed). The red dot is the classical LK critical inclination. Bottom: Deviation of the numerical solution to the polynomial least squares fit of order 40. The RMS value of the polynomial fit is $\Delta I_{RMS} = 0.01^\circ$.

In our investigation of the stability limit, the orbits have relative large separations, and thus large m values. The transformation from a to m is

$$\frac{a}{r_H} = 3^{1/3} m^{2/3} \quad (13)$$

The maximal eccentricity is modified from Eq. (6) to

$$e_{\max}(i, i_c(a)) = \sqrt{1 - \frac{\cos^2 i}{\cos^2 i_c(a)}} \quad (14)$$

and is different for each a . In this case, the critical radii depends on the maximal eccentricity (14) which in turn depends on the critical radius. In order to overcome this technical difficulty and find $i_c(a)$ in Eq. (14), we calculate first a temporary value of a , $r_{\text{temp}}(a, i) \equiv r_c(i) \cdot f(e_{\max}(a(m)))$ and look at $\Delta(a, i) \equiv r_{\text{temp}}(a, i) - a$. For small a , $\Delta(a, i) > 0$, or $a > r_{\text{temp}}(a, i)$ and the binary is stable. Only when $\Delta(a, i)$ changes sign, $a \geq r_{\text{temp}}(a, i)$, and the binary is unstable. Formally, the limiting radii

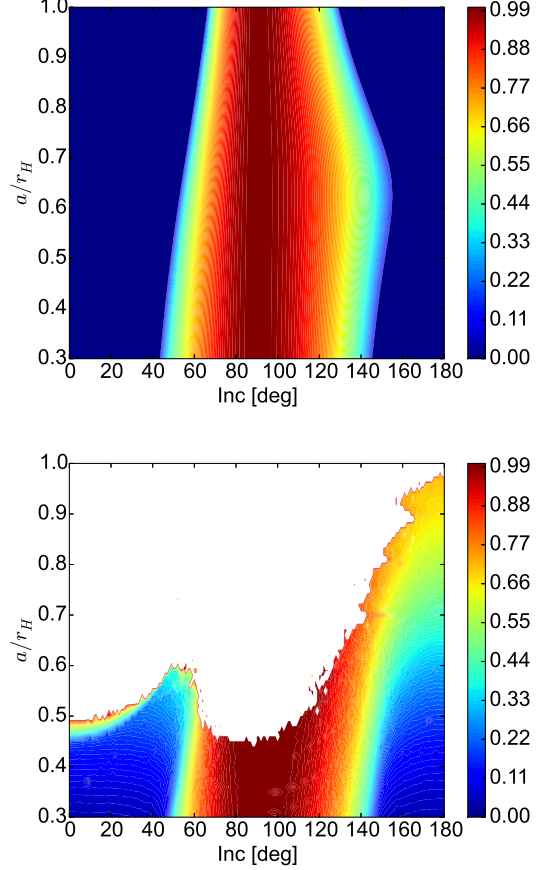


FIG. 4. Top: Maximal eccentricity of the LK mechanism combined with ER and LT (Eq. 14). Bottom: Simulated maximal eccentricity of the stable orbits.

is the minimum of the set

$$r_c(i) = \min_a \{\Delta(a, i) \leq 0\}. \quad (15)$$

Figure 4 shows the values of $e_{\max}(i, a)$ given in Eq. (14). e_{\max} is restored for the standard LK regime where $a \approx 0$. As a increases, the critical inclination shifts to the right to higher values. Thus, for prograde orbits the effective critical turn off point of the limiting radii curves will be shifted to the right. For retrograde orbits the shift will occur also to the right as long as $a/r_H \lesssim 0.6$. For higher values, the shift will be to the left. The bottom panel shows the resulted maximal eccentricity map of the numerical integration. Overall, the numerical results are in good agreement with the analytical expectations for orbits with $a \lesssim 0.6r_H$. For orbits with $a \gtrsim 0.6r_H$ the perturbations in eccentricity are large. Thus, the orbits are not simple Keplerian ellipses and the DA approximation breaks down.

We find the limiting radii in Eq. (15), and plot in Fig. 5 the additional corrections to r_c due to the evection terms. Due to the effective shift in i_c due to evection, most of the bump near $\sim 60^\circ$ is stable. The best fit

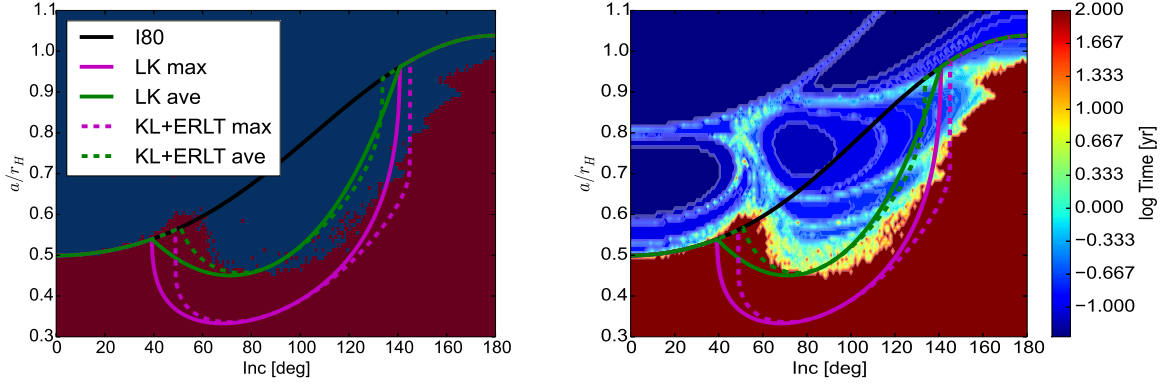


FIG. 5. The same as Fig 5, but with additional fits where e_{max} is determined both from LK and ER-LT terms (Eq. 14). The dashed green (magenta) lines are fits due to both LK and ER-LT for f_{ave} (f_{max}).

model is the the green dashed curve (LK and ERLT) is consistent with the numerical stability map for inclinations up to $i \lesssim 120^\circ$, where in the retrograde cases it fails to explain the stability map.

V. NUMERICAL PARAMETER SPACE EXPLORATION

A. Numerical set up

In order to test the critical stability radius, we have made numerical experiments, repeating, but also much extending the results of [6]. The planetary orbit is circular with semi-major axis $a_{out} = 1\text{AU}$. The mass ratios are $m_1/m_{out} = 10^{-6}$ and $m_2/m_1 = 10^{-2}$, thus the Hill radius $r_H = a_{out}(10^{-6}/3)^{1/3} \ll a_{out}$. We simulate a large grid of initial conditions of $a/r_H \in [0.3, 1.1]$ with steps $\delta a = 0.01r_H$ and $i \in [0^\circ, 180^\circ]$ with steps of $\delta i = 1^\circ$. The line of ascending node Ω , and the true mean M are zero. The argument of periape is ill defined for circular orbits. The total number of simulations is $N = 81 \times 181 = 14461$. The longest LK timescale is $T_{LK} = \sqrt{3}P_{out} \cdot 0.3^{-3/2} \approx 10.5$ years, while the orbits near the stability limit have much shorter timescale.

We conclude that a specific orbit is unstable in two cases: (i) If the binary eccentricity exceeds unity ($e > 1$). (ii) If the binary distance is larger than three times the Hill radius ($d_{bin} > 3r_H$). We use both criteria (i) and (ii) independently. Both criteria yield the same stability maps.

We simulate each initial condition for 100 years, at least one order of magnitude larger than the maximal LK timescale to capture the (quasi) secular effects. We use an N-body integrator with a shared but variable time step, using the Hermite 4th order integration scheme following Hut et al. [29].

The initial condition chosen to build the stability map of Fig. 5 limit the simulated parameter space to a degenerate subspace. The next sections explore more of the

parameter space and thus avoid any potential biases in the choice of the initial conditions. We will show that the overall morphology of the stability map is conserved and that the changes in the stability map are minor.

B. Circular grid

In order to explore more of the parameter space, we repeat the simulation, but now a resolution of $\delta i = 3^\circ$, and $a/r_H \in [0.3, 1]$ with $\delta a = 0.01r_H$. In addition, we randomly draw 100 initial conditions for the pair (M, Ω) for each value of a and i . Since the inner binary is circular, ω is not defined and not sampled. Thus, the total number of initial conditions is $N = 71 \times 61 \times 100 = 433,100$.

Fig. 6 shows the stability map of the circular grid. Top left (right) panel show the minimal (maximal) break up times of the stability map. The bottom left panel shows the percentage of the sampled stable orbits. We see that the stability limit has a chaotic boundary, and some width of $\sim 0.06r_H$. The orbits with zero mean anomaly are the most stable ones. Thus, we normalized the analytical fits by a fudge factor of 0.92, such that they lie roughly in the middle of the boundary for zero inclinations. Though the stability limit is lower, the general morphology of the stability boundary is preserved. For highly retrograde orbits, their stability limit has a cut off at $\sim 0.7 - 0.8r_H$, and some of the orbits above that limit are unstable. The chaotic boundary is also wider. The bottom right panel shows the comparison of the best analytical fit (green) to the maximal radius with 50% survival rate (blue). The shaded region is the 95% confidence level; above (below) the region, only 5% or less of the orbits are unstable (stable). We see that the analytical fit agrees well with the simulated stability grid, up to radii of $\lesssim 0.6r_H$ and inclination of $\lesssim 120^\circ$. In addition, the 95% confidence level zone is wider for retrograde orbits, thus indicating that the 95% confidence level is correlated to the chaotic boundary.

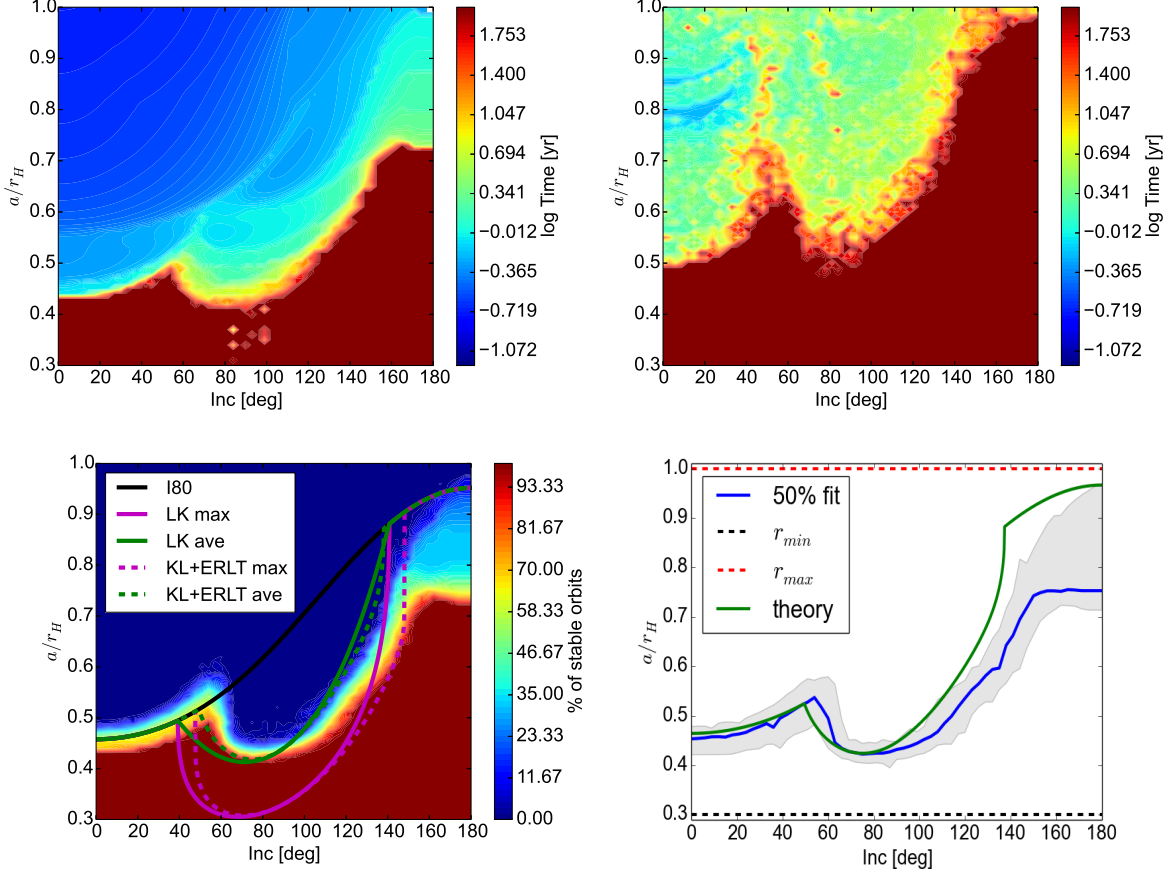


FIG. 6. Stability map of the circular grid, with 100 different realizations of each a and i . Top left: Time map of the minimal time of break up for the 100 realizations. Top right: Time map of the maximal time of break up for the 100 realizations. Bottom left: Percentage of the survived orbit with the analytical fits. Note different normalization of the analytical fits. Bottom right: Statistics of the destruction of the circular grid. The blue line corresponds to the survival of 50% of the orbits, the green line corresponds to the best fit analytical model (KL+ERLT ave). Grey region represents 95% confidence levels.

C. Eccentric grid

We repeat the simulations, but with $\delta i = 5^\circ$ and initial eccentricity $e = 0.1$. The range and spacing of the semi-major axis remains the same. We randomly draw 343 initial conditions for the triple (M, Ω, ω) for each value of a and i . The total number of initial conditions is $N = 71 \times 37 \times 343 = 901,061$.

Fig. 7 shows the stability map of the eccentric grid. Each panel is the same as in Fig 6. Besides slight changes in the width of the chaotic boundary and slightly lower stability limit in the retrograde case ($a \lesssim 0.7r_H$), the results are very similar to the circular grid in Fig. 6.

In both the circular and eccentric cases, the analytical fit is better since the angle of the “turn-off point” is slightly lower than $\sim 60^\circ$. We conclude that the initial simulations are biased toward the most stable orbits in the chaotic region, but still capture the general morphological picture. The circular and eccentric grid simulations help to overcome the bias and are a better fit for the analytical theory.

D. Polynomial fits

	a_0	a_1	a_2	a_3	a_4	q	i_{c1}	i_{c2}
p_{th}	4.22	-7.7	6.3	-2.34	0.339	0.721	0.86(49.4°)	2.41(138°)
p_{circ}	2.24	-2.03	0.6	0.68	-0.033	0.66	0.94(54°)	π
p_{ecc}	2.6	-2.9	1.3	-0.16	-0.0049	0.635	0.96(55°)	π

TABLE I. Polynomial fit coefficients for the polynomials defined by Eq. (16). p_{th} is the best analytic fit. p_{circ} is the polynomial fit of the numerical circular grid (Fig. 6). p_{ecc} is the polynomial fit of the numerical eccentric grid (Fig. 7).

For completeness, we provide 4th order polynomial fits for the best fit analytic model, the numerical results of the circular and eccentric grids. The polynomials are defined by

$$r_{th}(i)/r_H = q(g(i))^{-2/3} \times \begin{cases} p(i) & i_{c1} < i < i_{c2} \\ 1 & \text{else} \end{cases} \quad (16)$$

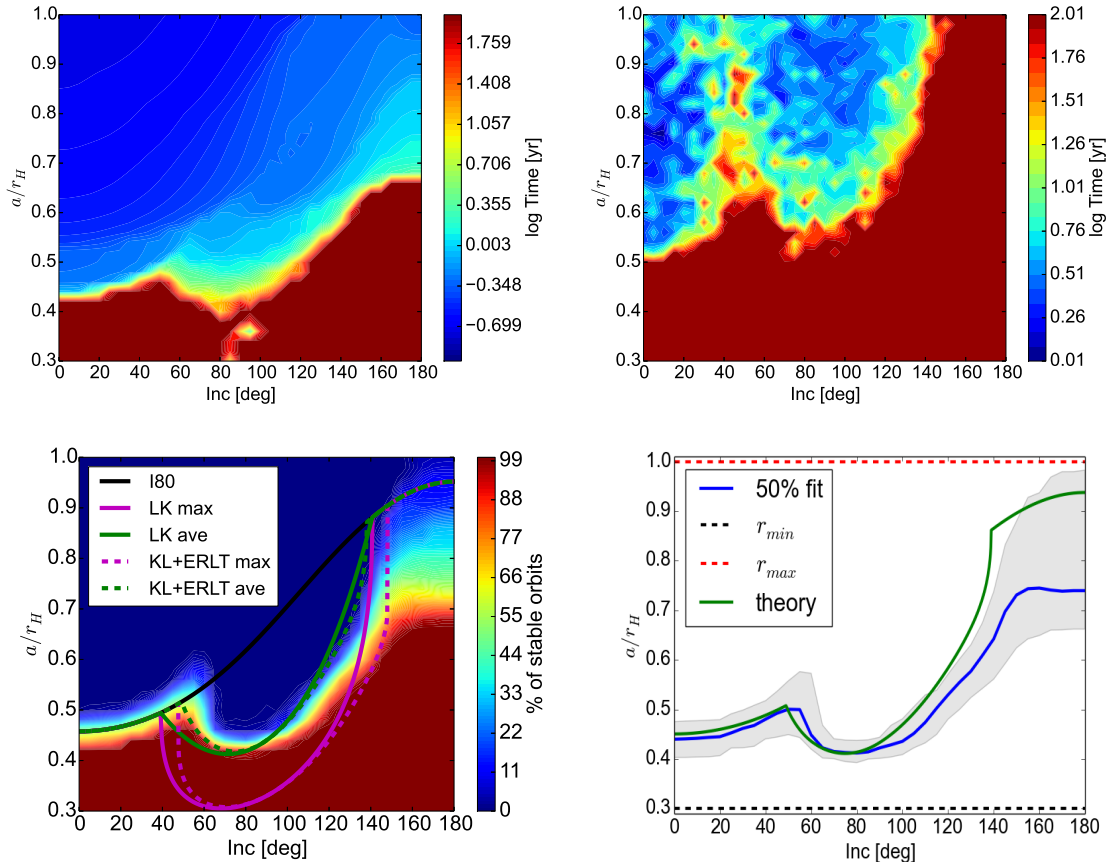


FIG. 7. The same as Fig. 6, but with initial eccentricity $e = 0.1$ and sampling of the argument of periapee ω .

where $g(i)$ is given in Eq. (3), q is the fudge factor and the polynomial $p(i) = \sum_0^4 a_j i^j$ is the best 4th order fit for the data. Coefficients and parameters of the fitting polynomials are summarized in table I.

VI. DISCUSSION AND SUMMARY

In this paper we have developed a generalized stability criteria for triple systems with arbitrary mutual inclination, and extended the Hill stability criteria to account for secular evolution and evection resonance. We used analytic arguments to derive the critical stability criteria and complemented them with extensive three-body simulations. Using these we also provided a convenient polynomial fitting formula for the stability criteria. The comparison between the analytical theory and numerical integrations is excellent, reproducing the morphology of the stability map, the amplitude and the "turn-off point" of the bowl-like-shaped stability region. In addition, the break up time-scales for highly inclined orbits is comparable with the quasi-secular Lidov-Kozai timescale, and much longer than the orbital period of the binary. This indicates that the mechanism that drives the highly inclined binaries to instability is indeed of secular nature,

rather than a dynamical Hill instability. In addition, the maximal eccentricity attained in the numerical integrations is compatible to the maximal eccentricity predicted by the analytic theory.

Though the analytic theory well describes prograde orbits, there is a significant discrepancy for retrograde orbits at high inclinations beyond $i \gtrsim 120^\circ$. At sufficiently large separations of $a/r_H \gtrsim 0.6$, the inner orbit of the triple is strongly perturbed by the outer companion and the orbit is no longer Keplerian. In addition, the timescales of the inner and outer binaries are comparable and the averaging method we use is no longer valid. The latter is evident in the eccentricity map of Fig. 4, where co-planar retrograde orbits, far from the Lidov-Kozai oscillation regime, also evolve into large eccentricities. For retrograde inclinations close to 180° , the analytic critical stability radius poorly describes the critical stability radius. Properly sampling the parameter space yields a chaotic transition boundary, which width increases for retrograde orbits. In this regime the critical stability strongly depends on the initial phase and orientation of the binary, such that the transition from stable to unstable orbits stretches to the range of $0.7 \lesssim a/r_H \lesssim 1$.

As an example we considered an application of our results for Solar System satellites. Our findings shed

light on the binomial inclination distributions of irregular satellites [23, 30]. In particular, the largest inclinations of prograde satellites is near $\sim 60^\circ$ (57° for Margaret, a moon of Uranus) and $\sim 130^\circ$ for retrograde satellites ($\sim 136.5^\circ$ for Neso, a moon of Neptune). At least for prograde satellites, the critical angle of $\sim 60^\circ$ is well explained by the evection corrections to the LK mechanism, as we show (see also [12]).

ACKNOWLEDGEMENTS

We would like to thank Scott Tremaine for useful discussions during the 33rd winter school of theoretical physics in Jerusalem. We acknowledge support from the Israel-US bi-national science foundation, BSF grant number 2012384, European union career integration grant “GRAND,” the Minerva center for life under extreme planetary conditions and the Israel science foundation excellence center I-CORE grant 1829/12.

Appendix: A. Analytical derivation of the minimal inclination

The implicit Eq. 3, is satisfied at least at one point $p_c = (i = i_{c,0}, m = 0)$. We assume that exists a local, unique solution $i \equiv i(m)$ in the vicinity of the point p_c .

Since $(d/di)h_0(i_c) \neq 0$, the assumptions of the implicit function theorem are satisfied, and a solution $i(m)$ can

be found locally as a power series in m . Deriving with respect to m Eq. 2 yields the first derivative:

$$i_{c,1} = \left. \frac{di(m)}{dm} \right|_{m=0} = - \left. \frac{h_1(i) \cos i}{h'_0(i)} \right|_{i=i_{c,0}} \quad (\text{A.1})$$

where $h'_0(i) = (d/di)h_0(i)$. Substitution at $i = i_c$ and $m = 0$ yields $i_1 = 53.5^\circ$.

Appendix: B. Coefficients of the least squares polynomial

1-10	11-20	21-30	31-40
53.2	$-1.015 \cdot 10^9$	$5.797 \cdot 10^{13}$	$3.202 \cdot 10^{15}$
-6.812	$-7.373 \cdot 10^9$	$2.478 \cdot 10^{13}$	$-6.76 \cdot 10^{14}$
213.348	$2.074 \cdot 10^{10}$	$-1.697 \cdot 10^{14}$	$-5.179 \cdot 10^{15}$
-3270.516	$8.133 \cdot 10^{10}$	$-1.925 \cdot 10^{13}$	$1.781 \cdot 10^{15}$
-3547.9	$-2.637 \cdot 10^{11}$	$3.021 \cdot 10^{14}$	$5.01 \cdot 10^{15}$
319181.311	$-6.128 \cdot 10^{11}$	$-5.828 \cdot 10^{13}$	$-2.182 \cdot 10^{15}$
-209339	$2.263 \cdot 10^{12}$	$-1.208 \cdot 10^{14}$	$-2.764 \cdot 10^{15}$
$-1.572 \cdot 10^7$	$3.2 \cdot 10^{12}$	$1.831 \cdot 10^{14}$	$1.399 \cdot 10^{15}$
$2.596 \cdot 10^7$	$-1.362 \cdot 10^{13}$	$-9.976 \cdot 10^{14}$	$6.739 \cdot 10^{14}$
$4.368 \cdot 10^8$	$-1.135 \cdot 10^{13}$	$-6.621 \cdot 10^{13}$	$-3.796 \cdot 10^{14}$

TABLE II. First 40 coefficients of the polynomial defined in Eq. (10).

-
- [1] R. A. Mardling IAUS **246**, 199-208 (2008).
 - [2] M. Holman, J. Touma, and S. Tremaine, Nature **386**, 254-256 (1997).
 - [3] K. A. Innanen, J. Q. Zheng, and S. Mikkola, and M. J. Vatonen, AJ **113**, 1915 (1997).
 - [4] V. Szebehely, *Theory of orbits. The restricted problem of three bodies* (1967).
 - [5] M. Henon, AA **9**, 24 (1970).
 - [6] D. P. Hamilton and J. A. Burns, Icarus **92**, 118 (1991).
 - [7] K. A. Innanen, AJ **84**, 960 (1979).
 - [8] K. A. Innanen, AJ **85**, 81 (1980).
 - [9] M. Valtonen and H. Karttunen, *The Three-Body Problem* (2006).
 - [10] M. L. Lidov, Planet. Space Sci. **9**, 719 (1962).
 - [11] Y. Kozai, AJ **67**, 591 (1962).
 - [12] M. Čuk and J. A. Burns, AJ **128**, 2518 (2004), astro-ph/0408119.
 - [13] L. Luo, B. Katz, and S. Dong, MNRAS **458**, 3060 (2016), 1601.04345.
 - [14] D. Nesvorný, J. L. A. Alvarelos, L. Dones, and H. F. Levison, AJ **126**, 398 (2003).
 - [15] J. Frouard, M. Fouchard, and A. Vienne, AA **515**, A54 (2010).
 - [16] B. Chauvineau and F. Mignard, Icarus **83** (1990).
 - [17] E. B. Ford, M. Havlickova, and F. A. Rasio, Icarus **150**, 303 (2001), astro-ph/0010178.
 - [18] Y. Lithwick and S. Naoz, ApJ **742**, 94 (2011), 1106.3329.
 - [19] S. Naoz, ArXiv e-prints (2016), 1601.07175.
 - [20] F. Antonini and H. B. Perets, ApJ **757**, 27 (2012), 1203.2938.
 - [21] H. B. Perets and S. Naoz, ApJL **699**, L17 (2009), 0809.2095.
 - [22] V. Carruba, J. A. Burns, P. D. Nicholson, and B. J. Gladman, Icarus **158**, 434 (2002).
 - [23] S. S. Sheppard, D. Jewitt, and J. Kleyna, AJ **129**, 518 (2005), astro-ph/0410059.
 - [24] D. Brouwer and G. M. Clemence, *Methods of celestial mechanics* (1961).
 - [25] P. Saha and S. Tremaine, Icarus **106**, 549 (1993).
 - [26] J. Frouard and T. Yokoyama, Celestial Mechanics and Dynamical Astronomy **115**, 59 (2013).
 - [27] F. Tisserand, *Traité de Mécanique Céleste, Tome III*. (Gauthier-Villars, Paris, 1894).
 - [28] A. Deprit, J. Henrard, and A. Rom, AJ **76**, 273 (1971).
 - [29] P. Hut, J. Makino, and S. McMillan, ApJL **443**, L93 (1995).
 - [30] J. N. Winn and D. C. Fabrycky, Annu. Rev. Astron. Astrophys. **53**, 409 (2015), 1410.4199.
 - [31] Note that ω is the argument of periapse of the satellite-planet system, not to be confused with the angular frequency of the satellite ω_1 . Similarly, Ω is the line of nodes of the satellite-planet system, not to be confused with the

angular frequency of the planet around the star Ω_{out} .

[32] The earliest list is in p. 233 of [27]. [25] use slightly different coefficients. [28] compares various coefficients derived

in the late 19th and early 20th century, namely Tisserand and Hill (see table III). [26] is a more modern work.